DATA-DRIVEN EXPLORATION OF MELODIC STRUCTURES IN HINDUSTANI MUSIC

Kaustuv Kanti Ganguli[1] **Sankalp Gulati**[2] **Xavier Serra**[2] **Preeti Rao**[1]

[1] Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai, India

[2] Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain

kaustuvkanti@ee.iitb.ac.in

ABSTRACT

Indian art music is quintessentially an improvisatory music form in which the line between 'fixed' and 'free' is extremely subtle. In a rāga performance, the melody is loosely constrained by the chosen composition but otherwise improvised in accordance with the rāga grammar. One of the melodic aspects that is governed by this grammar is the manner in which a melody evolves in time in the course of a performance. In this work, we aim to discover such implicit patterns or regularities present in the temporal evolution of vocal melodies of Hindustani music. We start by applying existing tools and techniques used in music information retrieval to a collection of concerts recordings of ālāp performances by renowned khayal vocal artists. We use svara-based and svara duration-based melodic features to study and quantify the manifestation of concepts such as vādi, samvādi, nyās and graha svara in the vocal performances. We show that the discovered patterns corroborate the musicological findings that describe the "unfolding" of a rāga in vocal performances of Hindustani music. The patterns discovered from the vocal melodies might help music students to learn improvisation and can complement the oral music pedagogy followed in this music tradition.

1. INTRODUCTION

Hindustani music is one of the two art music traditions of Indian art music [6], the other being Carnatic music [28]. Melodies in both these performance oriented music traditions are based on the framework of rāga [3]. Performance of a rāga in Indian art music (IAM) is primarily improvisatory in nature [26]. While some of these improvisations are based on a composed musical piece, Bandish, others are completely impromptu expositions of a rāga, Ālāp. Rāga acts as a grammar both in composition and in improvisation of melodies.

The rules of the rāga grammar are manifested at different time scales, at different levels of abstraction and demand a different degree of conformity. While some of the elements of the rāga grammar are explicit, others are implicit and may require years of musical training to grasp. A number of textbooks and musicological studies exist that describe different improvisatory aspects of melodies in IAM [1, 3, 4, 6, 10, 18, 26]. These works also attempt to uncover some of the implicit aspects of rāga grammar.

A majority of the studies mentioned above are musicological in nature. These typically involve either a thorough qualitative analysis of a handful of chosen musical excerpts or a compilation of expert domain knowledge. Though these studies often present interesting musical insights, there are several potential caveats in such works. Some of these caveats are summarized below:

- Small repertoire used in the studies challenge the generalizability of the proposed musical models
- Bias introduced due to the subjectivity in the analysis of musical excerpts
- Absence of concrete quantitative evidences supporting the arguments
- The kind of analysis that can be done (manually) is limited by human capabilities, limited memory (both short- and long-term)
- Difficulty in reproducibility of the results

In addition to the musicological works mentioned above, there are several studies that perform computational modeling of different melodic aspects in IAM. These studies address computational research tasks such as rāga recognition [5, 15], melodic similarity [11, 17, 20], discovery and search of melodic patterns [12, 16], segmentation of a musical piece [27] and identification of specific landmarks in melodies [14]. These approaches typically employ signal processing and machine learning methodologies to computationally model the relevant melodic aspect. These studies can provide a ground for developing tools and technologies needed to navigate and organize sizable audio collections of music, perform rāga-based search and retrieval from large audio archives and in several other pedagogical applications.

Several qualitative musicological works bring out new musical insights but are prone to criticism of not having supported their findings using a sizable corpus. Contrary to that, quantitative computational studies manage to scale to sizable data sets, but fall short of discovering novel musical insights. In the majority of cases, computational studies attempt to automate a task that is well known and is fairly easy for a musician to perform. There have been some studies that try to combine these two types of methodologies of working and corroborate several concepts in musical theories using computational approaches. In Chinese opera music, [22] has performed a comparison of the singing styles of two Jingju schools where the author exploit the potential of MIR techniques for supporting and enhancing musicological descriptions. Autrim [1] (Automated Transcription for Indian Music) has used MIR tools for visualization of Hindustani vocal concerts that created a great impact on music appreciation and pedagogy in IAM. We find that such literation pertaining to melodic structures in Indian art music is scarce.

In this paper, we perform a data-driven exploration of several melodic aspects of Hindustani music. The main objective of this paper is to use existing tools, techniques and methodologies in the domain of music information retrieval to support and enhance qualitative and descriptive musicological analysis of Hindustani music. For this we select five melodic aspects which are well described in musicological texts and are implicitly understood by musicians. Using computational approaches on a music collection that comprises representative recordings of Hindustani music we aim to study these implicit structures and quantify different melodic aspects related with them. In addition to corroborating existing musicological works, our findings are useful in several pedagogical applications. Furthermore, the proposed methodology can be used for analyzing artist or gharana-specific [2] melodic characteristics.

2. HINDUSTANI MUSIC CONCEPTS

Bor [3] remarks that rāga, although referred to as a concept, really escapes such categories as concept, type, model, pattern etc. Meer [26] comments that technically a rāga is a musical entity in which the intonation of svaras, as well as their relative duration and order, is defined. A rāga is characterized by a set of melodic features that include a set of notes ('svara'), the ascending and descending melodic progression ('ārōhana-avrōhana'), and a set of characteristic phrases ('pakad'). There are three broad categories of segments in the melody: stable svara regions, transitory regions and pauses. While svaras comprise certain hierarchical sub-categories like nyās, graha, amsa, vādi and samvādi, the transitions can also be grouped into a set of melodic ornamentation ('alankar') like meend, andolan, kan, khatka etc. [26]. The third important melodic event is the pause. Pauses carry much information about the phrases; in fact, a musician's skill lies in the judicious use of the pause [10]. We shall next go over these three broad aspects.

Many authors [3, 26] refer to the importance of certain svaras in a rāga. From the phrase outline we may filter certain svaras which can be used as rest, sonant or predominant; yet the individual function and importance of the svaras should not be stressed [26]. Nyās svara is defined as the resting svara, also referred to as 'pleasant pause' [7] or 'landmark' [14] in a melody. Vādi and samvādi are best understood in relation with melodic elaboration or vistār ('barhat'). Over the course of a barhat, artists make a particular svara 'shine' or 'sound'. There is often a corresponding svara which sustains the main svara and has a perfect fifth relation with it. The subtle inner quality of a rāga certainly lies in the duration of each svara in the context of the phraseology of the rāga. A vādi, therefore, is a tone that comes to shine, i.e., it becomes attractive, conspicuous and bright [26]. Another tone in the same rāga may become outstanding that provides an answer to the former tone. This second tone is the samvādi and should have a fifth relationship with the first tone. This relationship is of great importance.

A rāga is brought out through certain phrases that are linked to each other and in which the svaras have their proper relative duration. This does not mean that the duration, the recurrence and the order of svaras are fixed in a rāga; they are fixed only within a context [26]. The svaras form a scale, which may be different in ascending (ārōhana) and descending (avrōhana) phrases, while every svara has a limited possibility of duration depending on the phrase in which it occurs. Furthermore, the local order in which the svaras are used is rather fixed. The totality of these musical characteristics can best be laid down in a set of phrases ('calana') which is a gestalt that is immediately recognizable to the expert. In a rāga some phrases are obligatory while others are optional [26]. The former constitute the core of the rāga whereas the latter are elaborations or improvised versions. Specific ornamentation can add to the distinctive quality of the rāga [10].

There is a prescribed way in which a 'khayal' performance develops. The least mixed variety of khayal is that where an ālāp is sung, followed by a full sthāyi (first stanza of the composition) in madhya (medium) or drut (fast) laya (tempo), then layakari (rhythmic improvisation) and finally tān (fast vocal improvisation). When the barhat reaches the higher (octave) Sa (root svara of Hindustani music), the antara (second stanza of the composition) is sung. If the composition is not preceded by an ālāp, the full development of the rāga is done through barhat. The composition is based on the general lines of the rāga, the development of the rāga is again based on the model of the composition [26].

There are four main sources of a pause in a melody, these include: (i) gaps due to unvoiced consonants in the lyric, (ii) short breath pauses taken by the musician when out of breath, (iii) medium pauses where the musician shifts to a different phrase, and (iv) long pauses where the accompanying instruments improvise. Musically meaningful or musician-intended melodic chunks are delimited only by (iii) and (iv) [9].

Though Hindustani music is often regarded as improvisatory, the improvisation is structured. On a broader level there is a well defined structure within the space of

[1] https://autrimncpa.wordpress.com/

[2] Refers to a lineage or school of thought.

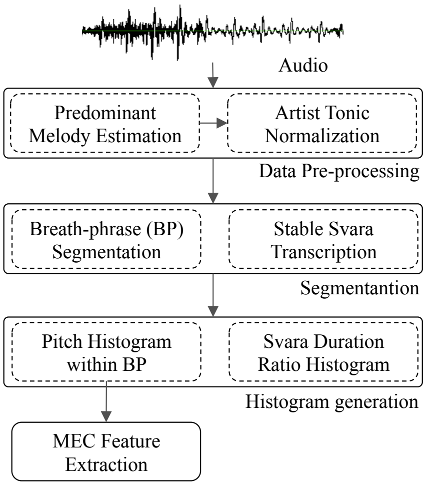

Figure 1. Block diagram for data processing

which an artist improvises following the rāga grammar. The overall skeleton of the melodies in Hindustani music can have defined structure at different levels. Some of these can be for the entire music tradition (for all rāgas and artists), some for specific rāgas, some of these can be gharana-specific, while some other artist specific. In this study, we aim to obtain overall structure in a melody and some rāga-specific patterns by the processing of audio concert recordings.

3. DATA PROCESSING

The block diagram for data processing is shown in Figure 1. It contains four main processing modules: preprocessing, segmentation, histogram generation and feature extraction. We describe these modules in detail in the subsequent sections.

3.1 Pre-processing

3.1.1 Predominant melody estimation

We start by estimating the pitch of the predominant melodic source in the audio signal and regard that as the representation of the melody. For predominant pitch estimation, we use the method proposed by Salamon and Gómez [23]. This method is reported to performed favorably in MIREX 2011 [3] on a variety of music genres, including IAM, and has been used in several other studies [8, 16]. We use the implementation of this algorithm as available in Essentia [2]. Essentia [4] is an open-source C++ library for audio analysis and content-based MIR. We use the default values of the parameters, except for the frame and hop sizes, which are set to 46 and 4.44 ms, respectively.

3.1.2 Tonic normalization

The base frequency chosen for a melody in a performance of IAM is the tonic pitch of the lead artist [13]. All other accompanying instruments are tuned with respect to this tonic pitch. Therefore, for analyzing features that are derived from the predominant melody across artists, the predominant melody should be normalized with respect to the tonic pitch. For this normalization we consider the tonic pitch ω as the reference frequency during the Hertz-to-cent-scale conversion, which is automatically identified using the multi-pitch approach proposed by Gulati et al. [13]. This approach is reported to obtain state-of-the-art results and has been successfully used elsewhere [12, 15].

3.2 Melodic segmentation

3.2.1 Breath-phrase segmentation

As described in Section 2 there are different types of unvoiced segments in the predominant melody. While some of these segments are musically a part of a melodic phrase (short-pauses), some others delineate consecutive melodic phrases. A distribution of the duration of all the unvoiced segments for the entire music collection revealed that their type can be identified based on the duration of the pause. For identifying intended breath pauses that separate melodic phrases we empirically set the duration threshold to be 500 ms. The duration of the intra-phrase pauses is considerably smaller than this threshold. All the intra-phrase breath pauses (i.e., with duration smaller than 500 ms) are interpolated using a cubic spline curve. We shall refer to a breath-phrase as BP hereafter.

3.2.2 Stable svara transcription

In Indian art music, written notation has a purely prescriptive role. Transcribing the melody of a musical performance into a written notation is a challenging task and requires an in-depth knowledge of the rāga [21, 29]. In this study we consider a simple melodic transcription that retains only the stable svara regions of a pitch contour and discards the transitory pitch regions. We first detect all the valid svaras and their precise frequencies used in a melody by computing a pitch histogram. Subsequently, we segment the stable svara regions by identifying the fragments of pitch contour that are within 35 cents [20] of the svara frequencies. Next, we filter out the svara fragments that are smaller than 250 ms in duration, as they are too short to be considered as perceptually meaningful held svaras [19]. This leaves a string of fragments each labeled by a svara. Fragments with the same svara value that are separated by gaps less than 100 ms are merged [12]. The resulting symbol sequence thus comprises a tuple of svara name and duration.

3.3 Histogram generation

3.3.1 Pitch histogram of breath-phrases

We compute the histogram of the transcribed svaras corresponding to each BP. Figure 2 shows the pitch histogram for each BP of the concert of rāga Todi sung by Ajoy Chakrabarty. . The 12th bin along the y-axis corresponds to the tonic svara Sa (0 cents), 24th for its octave (1200 cents).

[3] http://www.music-ir.org/mirex/wiki/2011
[4] https://github.com/MTG/essentia

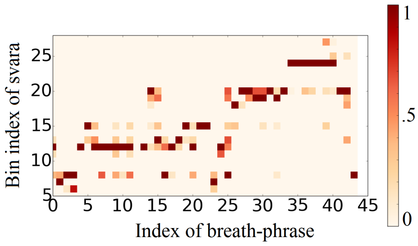

Figure 2. Pitch histogram of svaras for each breath-phrase.

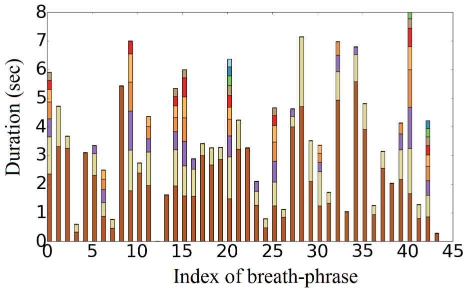

Figure 3. Bar graph of svara duration stacked in sorted manner for each breath-phrase. We observe that breath-phrases often comprise one long nyas svara and several other svaras of less duration.

3.3.2 Svara duration distribution

We consider the distribution of the svara duration for each BP. Figure 3 shows a stacked bar graph of the sorted durations of the svaras for each BP for the same concert. We observe that the cumulative duration of the transcribed svaras range from approximately 1 to 8 seconds. An important point to note here is that there is a difference between the absolute duration of a BP and the height of the stacked bar (in Figure 3). This is caused by the transient pitch segments that we ignored in our representation. Readers must note that the stable svara transcription, therefore, has an implication for the further analyses.

3.4 Post-processing

To capture the changes in the svara pattern at a broader time-scale, we time-average the pitch histogram across ten BPs with a hop of one BP. This is followed by tracking of the most salient bin across the smoothened histogram . Finally, the obtained contour is further median filtered with one preceding and succeeding BP. We refer to this contour as the *evolution contour* (hereafter EC). Figure 4 shows the time-averaged histogram superimposed with the EC for the same concert.

3.5 MEC feature extraction

We would like to compare the ECs of different concerts to explore whether there is any temporal trend that holds across the collection. To normalize the time-scale and pitch range of the EC, we normalize each EC within a unit

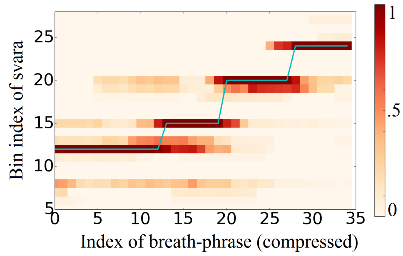

Figure 4. Time-averaged pitch histogram superimposed with the evolution contour.

range in both temporal and pitch dimensions. Thus a *modified evolution contour* (hereafter MEC) is obtained as:

$$MEC = \frac{EC - \min(EC)}{\max(EC) - \min(EC)} \tag{1}$$

with 100 samples between [0,1] .

We extract a collection of heuristic features (slope-based, duration-based, jump-based and level-based) from the MEC. A few important features are: slope between the MEC value of 0^{th} frame and the first frame where $MEC = 1$ (referred to as Slp) , proportion of duration spent on each svara (referred to as Pro), centroid (considering salience of the bins as the weights in the centroid computation) of each svara (referred to as Cen) , starting and ending svaras, (second) longest svara and proportion of its duration, magnitude of lowest/highest jumps between consecutive levels etc.

4. MUSIC COLLECTION AND ANNOTATIONS

The music collection used in this study is taken from the Hindustani music dataset (HMD) compiled as a part of the CompMusic project [24, 25] (130 hours of commercially available audio recordings stored as 160 kbps mp3 stereo audio files). All the editorial metadata for each audio recording is publicly available on an open-source metadata repository called MusicBrainz [5]. The selected music material in our collection is diverse in terms of the number of artists (40), recordings (mostly live concerts of both male and female renowned musicians from the last 6 decades) and the number of unique compositions (67). In these terms, it can therefore be regarded as a representative subset of real-world collections. Our collection includes a total of 75 concerts from 10 widely used rāgas (8 pieces per rāga on an average) that are diverse both in terms of the number of svaras and their pitch-classes (svarasthānās). All the concerts belong to either madhya or drut laya (and non-metered ālāp). The pitch range of the recordings spans approximately two octaves (middle octave and half of the lower/upper octave). All of the concerts comprise elaborations based on a bandish.

The scope of the study is limited to only the ālāp and vistār (barhat) [3, 4, 6, 18, 26] sections of the concert. Almost all of the concerts continue to subsequent fast improvisatory section (tān) after rendering the vistār. The

[5] https://musicbrainz.org/

melodic cue where the antara ends and the rhythmic cue where there is a slight increase in tempo just after, is quite universal and musically unambiguous. For reference, please observe Section 3.1 in [30]. We employ a performing Hindustani musician (trained over 20 years) to annotate the time-stamps where the vistār ends. As the said annotation can be considered an obvious one (there is a less chance of getting subjective biases), we limit the manual annotation to one subject only. After cutting the concerts to the musician-annotated time-stamp, the average duration per concert is 16 minutes making a total of 20 hours of data.

5. ANALYSIS AND DISCUSSION

We choose certain music concepts which are widely discussed among musicians and musicologists, for which there has not yet been an objective way of interpreting them from the audio. We cite the concepts (or knowledge-hypotheses, referred to as 'K') and discuss how a data-driven approach can help us validate them.

5.1 K1: Evolution of melody in time

As discussed in Section 2, the barhat of a rāga performance refers to the gradual "unfolding" of a rāga by building on the svaras with a progression in a broad time-scale. But it is not very clearly illustrated in the musicology literature what the precise duration is spent on each svara in course of this progression. Figure 5 shows the MECs of 37 (50% randomly chosen from our collection) concerts. We observe that the MECs, in general, start from a lower svara and gradually reach the highest svara in a step-like manner. The slope Slp of MEC, is quite consistent (mean = 1.3, standard deviation = 0.34) over the whole corpus. This gives an important insight that irrespective of the rāga and concert-duration, artists take the same time to explore the melody and hit the highest svara. This also reinforces the nature of the time-scaling of a performance: for either a short 20 minute- or a long 50 minute-concert, the melodic organization bases more on relative and not absolute time. We also observe a sharp fall of the MEC at the end of the many concerts, this reflects how artists come down to a lower svara to mark an end to the vistār (this coincides with the musician's annotation). This phenomenon has a high resemblance with the time evolution of melody in course of the vistār, as shown in Figure 11 in [30].

5.2 K2: Transitional characteristics of nyās svaras

Rāga guidelines mention about allowed svara sequences within a melodic phrase but it would be interesting to see if artists maintain any specific order in choosing the nyās svara across BPs or take liberty to emphasize any other svara. This is to be captured from the granularity of BPs and not in the time-averaged MEC. We generate a svara-transition matrix and populate it with a uniform weight for all transitions of the salient bins across BPs. Figure 6 shows the salient svara-transition matrix where the diagonal elements refer to self transitions. As indicative from

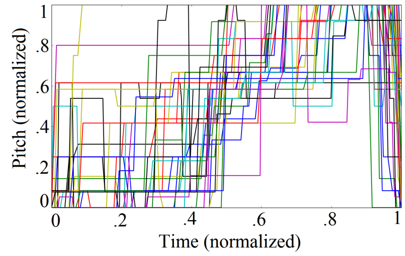

Figure 5. Modified evolution contours for 37 concerts in our music collection.

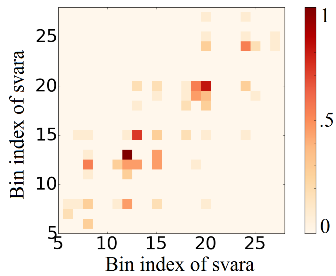

Figure 6. Svara-transition matrix of salient svaras of each breath-phrase. Intensity of each bin is proportional to the number of transitions taken from the svara of bin index on x-axis to the svara of bin index on y-axis.

wide steps of the MECs, there are quite a few self transitions but to our interest the salient transitions across BPs also follow a pattern alike the allowed svara-transitions within a melodic phrase. This is not a trivial event. We compute a feature to measure the steadiness quotient Stq of the transition matrix, defined as the ratio of the trace of the svara-transition matrix to the sum of all bins. We observe a very low standard deviation (0.23) across our music collection which conforms to the fact that artists 'focus' on a nyās svara for consecutive BPs to establish that svara.

5.3 K3: Relationship between functional roles of svaras and their duration in melody

We discussed about functional roles of vādi/samvādi svaras, but it is not explicitly known how their 'prominence' is defined. Earlier work [5] use histogram and show that they are one of the most used tonal pitches. But it is not evident from a pitch histogram whether the peak heights are contributed by a large number of 'short' svara segments or a fewer 'long' svara segments. Figure 7 shows the mean (left) and standard deviation (right) of all svaras (octave folded) for each svara along x-axis being the salient svara in a BP. We observe that the role of each svara is defined in terms of their duration in context of a nyās svara. This conforms with the concepts discussed in Section 2. This also reconfirms the well-defined structure of the melodic improvisation that any svara cannot be stretched arbitrarily long, the nyās svara of the BP and the phrase itself decides how much variance all other svaras can incorporate. This also brings out a question whether there is any special func-

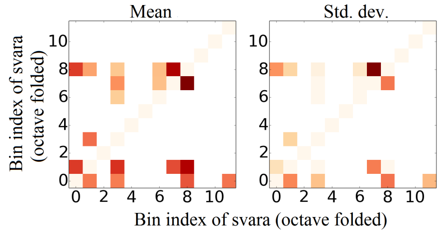

Figure 7. Mean (left) and standard deviation (right) of all svaras (octave folded) for each svara along x-axis being the salient svara in a breath-phrase.

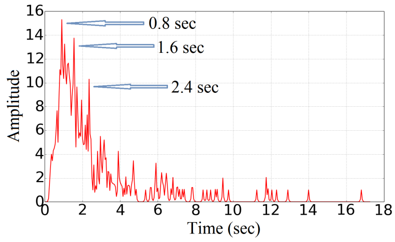

Figure 8. Ratio of inter-onset-interval of salient svaras across breath-phrases. We see a tatum pulse (peak) at 0.8 seconds and its harmonics.

tional role of vādi/samvādi svara in terms of their duration in a specific BP. One observation is that the vādi/samvādi svara, while a salient svara in the respective BPs, constrain the duration of all other svaras, making its relative prominence higher.

5.4 K4: Duration and position of svaras in melody

In music theory, vādi and samvādi svaras are among the concepts which are often discussed. But we do not have an objective measure to observe how these svaras are different from other svaras in a rāga. It is also interesting to know whether the vādi or samvādi svara always takes a focal role irrespective of their location in a BP and the overall position in a performance.

Position in melody: An important feature of the MEC is the Cen. We observe that the Cen is a rāga dependent feature. E.g., an uttaranga vādi rāga would have its vādi centroid in the latter half of a concert. This is supportive of the fact that the vādi is explored in due course of melodic improvisation adhering to the trend observed as in Section 5.1. The musicological hypothesis that these are the focal svaras of a rāga does not necessarily imply that these svaras are explored from the very beginning of the concert. Rather the performance starts from a lower svara (graha svara) and reaches the vādi in course of the gradual development of the melody.

Duration in melody: We compute the average duration of all salient svaras per BP in two groups of svaras: (i) group 1: vādi/samvādi, and (ii) group 2: rest. It is observed on the whole corpus that Pro of group 1 is higher than the group 2 for all rāgas. This reinforces the fact the term 'focus' or 'shine' (that qualifies vādi) is manifested in the temporal dimension. This also brings out a question whether we can predict the vādi/samvādi of a rāga from the melody by data-driven features. From the overall pitch histogram it is difficult to infer, but from our designed features, we observe an accuracy of 83% while predicting the vādi/samvādi of a given raga.

5.5 K5: Presence of possible pulsation in melody

There has been a discussion among musicians and musicologists whether there exists a pulsation in the melody of an ālāp in Hindustani music. Musicians agree there is an implicit pulsation present, but quantification is left to subjects. At the same time, the subjective bias only results in an octave difference, i.e., there is a harmonic relation among the pace in which the subjects tap to the melody. We propose a measure, through our data-driven approach, to estimate a possible metric for the pulsation. We assume that the pulse obtained from the melody would correlate to the percussive pulsation. We compute the ratio of inter-onset-interval of the most salient svaras across BPs. Figure 8 shows a pulsation at 0.8 seconds and its harmonics which correspond to 75 beats per minute (bpm) and the percussive tempo of the concert is approximately 40 bpm. The noise in the estimate may also follow from a few short BPs (e.g., BP index 3, 7 etc.) as observed in Figure 3. This measure, therefore, needs further investigation before we generalize over the corpus.

6. CONCLUSION AND FUTURE WORK

In this paper we performed a data driven exploration of implicit structures in melodies of Hindustani music. We outlined the motivation and relevance of computational approaches for quantitatively studying the underlying musical concepts. We computed musically relevant and easy-to-interpret acoustic features such as svara frequency and svara duration histogram. For computing these features we primarily used existing tools and techniques often used in information retrieval of Indian art music. We performed a quantitative analysis of 75 music concerts in Hindustani music in 10 different rāgas. With that we showed how the musical concepts are manifested in real-world performances and experimented with several ways to quantify them. With this we also corroborate some of the interesting music concepts and discover implicit relationships between svaras and duration in the temporal evolution of a rāga performance. In the future, one possible research direction would be to use these findings for characterizing artist-specific aspects and highlighting different nuances across gharanas in Hindustani music.

7. ACKNOWLEDGEMENT

This work received partial funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement 267583 (CompMusic). Part of the work was supported by Bharti Centre for Communication in IIT Bombay.

8. REFERENCES

[1] S. Bagchee. *Nād: Understanding Rāga Music*. Business Publications Inc, 1998.

[2] D. Bogdanov, N. Wack, E. Gómez, S. Gulati, P. Herrera, O. Mayor, G. Roma, J. Salamon, J. Zapata, and X. Serra. Essentia: An audio analysis library for Music Information Retrieval. In *Proc. of Int. Soc. for Music Information Retrieval (ISMIR)*, pages 493–498, 2013.

[3] J. Bor, S. Rao, W. van der Meer, and J. Harvey. *The Raga Guide: A survey of 74 Hindustani ragas*. Nimbus Records with Rotterdam Conservatory of Music, 1999.

[4] A. Chakrabarty. *Shrutinandan: Towards Universal Music*. Macmillan, 2002.

[5] P. Chordia and S. Şentürk. Joint recognition of raag and tonic in north Indian music. *Computer Music Journal*, 37(3):82–98, 2013.

[6] A. Danielou. *The ragas of Northern Indian music*. Munshiram Manoharlal Publishers, New Delhi, 2010.

[7] A. K. Dey. *Nyasa in raga: The pleasant pause in Hindustani music*. Kanishka Publishers, 2008.

[8] S. Dutta and H. A. Murthy. Discovering typical motifs of a raga from one-liners of songs in Carnatic music. In *Int. Soc. for Music Information Retrieval (ISMIR)*, pages 397–402, Taipei, Taiwan, 2014.

[9] K. K. Ganguli. Guruji Padmashree Pt. Ajoy Chakrabarty: As I have seen Him. *Samakalika Sangeetham*, 3(2):93–100, October 2012.

[10] K. K. Ganguli. How do we 'See' & 'Say' a raga: A Perspective Canvas. *Samakalika Sangeetham*, 4(2):112–119, October 2013.

[11] K. K. Ganguli and P. Rao. Discrimination of melodic patterns in Indian classical music. In *Proc. of National Conference on Communications (NCC)*, February 2015.

[12] K. K. Ganguli, A. Rastogi, V. Pandit, P. Kantan, and P. Rao. Efficient melodic query based audio search for Hindustani vocal compositions. In *Proc. of Int. Soc. for Music Information Retrieval (ISMIR)*, pages 591–597, October 2015. Malaga, Spain.

[13] S. Gulati, A. Bellur, J. Salamon, H. G. Ranjani, V. Ishwar, H. A. Murthy, and X. Serra. Automatic tonic identification in Indian art music: approaches and evaluation. *Journal of New Music Research (JNMR)*, 43(1):55–73, 2014.

[14] S. Gulati, J. Serrà, K. K. Ganguli, and X. Serra. Landmark detection in Hindustani music melodies. In *International Computer Music Conference, Sound and Music Computing Conference*, pages 1062–1068, Athens, Greece, 2014.

[15] S. Gulati, J. Serra, V. Ishwar, S. Senturk, and X. Serra. Phrase-based rāga recognition using vector space modeling. In *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 66–70, Shanghai, China, 2016.

[16] S. Gulati, J. Serrà, V. Ishwar, and X. Serra. Mining melodic patterns in large audio collections of Indian art music. In *Int. Conf. on Signal Image Technology & Internet Based Systems (SITIS-MIRA)*, pages 264–271, Morocco, 2014.

[17] S. Gulati, J. Serrà, and X. Serra. Improving melodic similarity in Indian art music using culture-specific melodic characteristics. In *Proc. of Int. Soc. for Music Information Retrieval (ISMIR)*, pages 680–686, Malaga, Spain, 2015.

[18] N. A. Jairazbhoy. *The Rags of North Indian Music: Their Structure & Evolution*. Popular Prakashan, second edition, 2011.

[19] P. Rao, J. C. Ross, and K. K. Ganguli. Distinguishing raga-specific intonation of phrases with audio analysis. *Ninaad*, 26-27(1):59–68, December 2013.

[20] P. Rao, J. C. Ross, K. K. Ganguli, V. Pandit, V. Ishwar, A. Bellur, and H. A. Murthy. Classification of melodic motifs in raga music with time-series matching. *Journal of New Music Research (JNMR)*, 43(1):115–131, April 2014.

[21] S. Rao and P. Rao. An overview of Hindustani music in the context of Computational Musicology. *Journal of New Music Research (JNMR)*, 43(1):24–33, April 2014.

[22] R. C. Repetto, R. Gong, N. Kroher, and X. Serra. Comparison of the singing style of two jingju schools. In *International Society for Music Information Retrieval (ISMIR)*, pages 507–513, Malaga, Spain, October 2015.

[23] J. Salamon and E. Gómez. Melody extraction from polyphonic music signals using pitch contour characteristics. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(6):1759–1770, 2012.

[24] X. Serra. A multicultural approach to music information research. In *Proc. of Int. Conf. on Music Information Retrieval (ISMIR)*, pages 151–156, 2011.

[25] X. Serra. Creating research corpora for the computational study of music: the case of the Compmusic project. In *Proc. of the 53rd AES Int. Conf. on Semantic Audio*, London, 2014.

[26] W. van der Meer. *Hindustani music in the 20th century*. Martin US Nijhoff Publishers, 1980.

[27] P. Verma, T. P. Vinutha, P. Pandit, and P. Rao. Structural segmentation of hindustani concert audio with posterior features. In *Int.Conf. on Acoustics Speech and Signal Processing (ICASSP)*, pages 136–140. IEEE, 2015.

[28] T. Viswanathan and M. H. Allen. *Music in South India*. Oxford University Press, 2004.

[29] R. Widdess. Involving the performers in transcription and analysis: a collaborative approach to Dhrupad. *Ethnomusicology*, 38(1):59–79, 1994.

[30] R. Widdess. Dynamics of melodic discourse in indian music: Budhaditya mukherjee's ālāp in rāg Pūriyā-Kalyān. pages 187–224, 2011.